/ 13

Noncommuting Observables in Quantum Detection and Estimation Theory

Carl W. Helstrom*

Department of Applied Physics and Information Science

University of California, San Diego

La Jolla, California 92037

(NASA-CR-122364) NONCOMMUNTING OBSERVABLES IN QUANTUM DETECTION AND ESTIMATION THEORY C.W. Helstrom (California Univ.) 10 Dec. 1971 8 p CSCL 20H

N72-21675

Unclas G3/24 23634

G 3

Abstract

In quantum detection theory the optimum detection operators must commute; admitting simultaneous approximate measurement of noncommuting observables cannot yield a lower Bayes cost. The lower bounds on mean square errors of parameter estimates predicted by the quantum-mechanical Cramér-Rao inequality can also not be reduced by such means.

* This research was supported by Grant NGL 05-009-079 from the National Aeronautics and Space Administration.

PAGES - 8
CATEGORY - 24

Quantum detection and estimation theory has been developed within the conventional framework of quantum mechanics, one of the principal tenets of which is that only observables associated with commuting operators can be simultaneously measured on the same system.[1-3] It has been suggested that this formulation is too restrictive, that noncommuting operators can be at least approximately measured on the same system, and that to include this possibility may permit more effective detection, as measured by a lower average Bayes cost.[4,5] We wish to show that no such improvement can be expected.

The simultaneous measurement of noncommuting observables has been treated by Gordon and Louisell.[6] In order to approximately measure certain such observables on a quantum-mechanical system S, it is made to interact for a time with a second system A, termed the apparatus. It was shown that a suitably defined ideal measurement yielding approximate values of the noncommuting observables can be based on the outcome of measurements of commuting observables on the apparatus A, or more generally on both S and A. What we must therefore do is apply quantum detection theory--with its restriction to commuting observables--to the combined system S + A.

Suppose we are to decide among M hypotheses $H_1, H_2, \ldots, H_M$. Under hypothesis H_j the density operator for the combined system at time t is $\rho_j^{S+A}(t)$ in the Schrödinger picture. If at an earlier time t_0 the density operator is $\rho_j^{S+A}(t_0)$, the two operators are related by[7]

$$\rho_j^{S+A}(t) = U(t, t_0)\, \rho_j^{S+A}(t_0)\, U^+(t, t_0), \tag{1}$$

with

$$U(t, t_0) = \exp\left[-i \int_{t_0}^{t} H dt'/\hbar\right], \tag{2}$$

where H is the Hamiltonian operator for the combined system S + A and $\hbar$ is

Planck's constant $h/2\pi$. The operator U is unitary; that is, with U^+ its Hermitian adjoint, UU^+ equals the identity operator $\underset{\sim}{1}$.

Let $\{\Pi_j\}$ be a set of commuting projection operators forming an M-fold resolution of the identity,

$$\sum_{j=1}^{M} \Pi_j = \underset{\sim}{1}. \qquad (3)$$

On the combination S + A we are to measure these M projection operators at time t, and if the k-th yields the value 1, hypothesis H_k is selected as true.[1] The average cost is then

$$\bar{C} = \sum_{i=1}^{M}\sum_{j=1}^{M} \zeta_j \, C_{ij} \, \mathrm{Tr}[\rho_j^{S+A}(t) \, \Pi_i], \qquad (4)$$

where ζ_j is the prior probability of hypothesis H_j and C_{ij} is the cost of choosing H_i when H_j is true. Let $\{\Pi_j(t)\}$ be the projection operators that minimize $\bar{C}$ when the system S + A is observed at time t; we call these optimum. Then by (1) the operators

$$\Pi_j(t_0) = U^+(t, t_0) \, \Pi_j(t) \, U(t, t_0) \qquad (5)$$

will minimize $\bar{C}$ when S + A is observed at time t_0. Because of the unitarity of $U(t, t_0)$, the set $\{\Pi_j(t_0)\}$ also forms an M-fold resolution of the identity into commuting projection operators, and the $\Pi_j(t_0)$ are optimum at time t_0. Since the minimization is carried out over all possible M-fold resolutions of identity, the minimum Bayes cost $\bar{C}_{min}$ must be independent of the observation time t.

Now let us roll time back to an epoch t_0 before the system S has come into contact with the apparatus A. In the Schrödinger picture this amounts to applying the inverse unitary transformation $U^+(t, t_0)$ to the state vectors of the combined system S + A. Because S and A are independent at this time t_0, the density operators ρ_j^{S+A} must now have the factored form

$$\rho_j^{S+A}(t_0) = \rho_j^{S}(t_0)\, \rho^{A}(t_0), \quad j = 1, 2, \ldots M. \tag{6}$$

Furthermore, as the apparatus A before the interaction has no information about which hypothesis is true, $\rho^A(t_0)$ in (6) must be independent of j. The Bayes cost is now

$$\bar{C} = \mathrm{Tr}\left[\rho^{A}(t_0) \sum_{i=1}^{M} \sum_{j=1}^{M} \zeta_j\, C_{ij}\, \rho_j^{S}(t_0)\, \Pi_i\right]. \tag{7}$$

Since S and A are completely uncoupled, and the state of A is independent of which hypothesis H_j is true, there is nothing to be gained by observing A. The optimum projection operators $\Pi_i(t_0)$ factor as $\Pi_i^S(t_0)\, \underset{\sim}{1}^A$, where $\underset{\sim}{1}^A$ is the identity operator for the apparatus A, and the set $\{\Pi_j^S(t_0)\}$ forms an M-fold resolution of the identity $\underset{\sim}{1}^S$ for the system S, minimizing the Bayes cost

$$\bar{C}_S = \sum_{i=1}^{M} \sum_{j=1}^{M} \zeta_j\, C_{ij}\, \mathrm{Tr}[\rho_j^{S}(t_0)\, \Pi_i]. \tag{8}$$

Since $\mathrm{Tr}[\rho^A\, \underset{\sim}{1}^A] = 1$, the minimum value of $\bar{C}_S$ is also the minimum value of $\bar{C}$ in (7) and equals the time-independent minimum Bayes cost $\bar{C}_{min}$. The decision among the M hypotheses made at time t_0 is based entirely on the measurement of commuting observables on system S.

Similar considerations apply to estimating the m parameters $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ of the density operator $\rho^S(\underset{\sim}{\theta})$ of a quantum-mechanical system S. A version of the Cramér-Rao inequality sets lower bounds to mean square errors of unbiased estimates of $\theta_1, \theta_2, \ldots, \theta_m$.[3] Let X_j be an operator whose measurement on S yields an unbiased estimate $\hat{\theta}_j$ of the j-th parameter; $\hat{\theta}_j$ must be an eigenvalue of X_j. Although in order to be measured simultaneously on the same system the operators X_j must commute, the analysis leading to the lower bounds given in [3] does not require commutativity of the operators X_j.

For the class of *commuting* operators yielding unbiased estimates of the parameters $\underset{\sim}{\theta}$ there will exist lower bounds on the mean square errors, and those will be greater than or at least equal to the bounds derived in [3].

Again including the possibility of measuring noncommuting observables cannot lead to lower bounds smaller than those in [3]. In order to measure such operators even approximately, a measuring apparatus A must be allowed to interact with the system S, and according to Gordon and Louisell's treatment of the process, commuting operators will at the end be measured on the combined system S + A.[6] In the Schrödinger picture the density operator $\rho^{S+A}(\underset{\sim}{\theta}, t)$ for S + A will have a time dependence similar to that in (1).

Referring to (7) of [3] we see that the symmetrized logarithmic derivatives (SLD) $L_j(t)$ appropriate for determining the Cramér-Rao lower bounds when the measurements of X_j are made at time t are related to those appropriate for measurements made at t_0 by

$$L_j(t) = U(t, t_0)\, L_j(t_0)\, U^+(t, t_0). \qquad (9)$$

Then (13) of [3] shows that the matrix $\underset{\sim}{A}$ that sets the lower bounds is independent of the time t of observation, again because of the unitarity of the operator $U(t, t_0)$.

Once more we move back to an epoch t_0 before the system and the apparatus have interacted. The density operator $\rho^{S+A}(\underset{\sim}{\theta}, t_0)$ factors as $\rho^S(\underset{\sim}{\theta}, t_0)\, \rho^A(t_0)$, where the density operator $\rho^A(t_0)$ of the apparatus A is independent of the estimanda $\underset{\sim}{\theta}$. The SLD operators for calculating the lower bounds are now the solutions of the operator equations

$$\partial\rho^S(\underset{\sim}{\theta}, t_0)/\partial\theta_n = \frac{1}{2}\, [\rho^S(\underset{\sim}{\theta}, t_0)\, L_n{}^S + L_n{}^S\, \rho^S(\underset{\sim}{\theta}, t_0)], \qquad (10)$$

and they act only on system S, commuting with ρ^A and all other operators on the apparatus A. When taking the trace over the states of A to form the elements of the matrix $\underset{\sim}{A}$, the density operator ρ^A is replaced by 1, and the lower bounds depend only on $\rho^S(\underset{\sim}{\theta}, t_0)$. Thus the apparatus A cannot help estimate the parameters $\underset{\sim}{\theta}$ of S with smaller mean square errors than the lower bounds calculated by the quantum-mechanical Cramér-Rao inequality as applied to the density operator of system S alone.

In [3, p. 238] lower bounds were calculated for unbiased estimates $\hat{m}_x$ and $\hat{m}_y$ of the components of the complex amplitude $\mu = m_x + im_y$ of a simple harmonic oscillator, which might represent a mode of the field in an ideal receiver in the presence of thermal noise. Those bounds are

$$\mathrm{Var}\ \hat{m}_x \geq \frac{1}{2}(N + \frac{1}{2}), \qquad \mathrm{Var}\ \hat{m}_y \geq \frac{1}{2}(N + \frac{1}{2}),$$

where N is the mean number of noise photons. The noncommutativity of the SLD's L_x and L_y used to derive these bounds does not invalidate them. It can be shown that if the mode is coupled with an ideal amplifier whose gain is high enough to raise the oscillator variables to the classical domain where they commute, error variances

$$\mathrm{Var}\ \hat{m}_x = \mathrm{Var}\ \hat{m}_y = \frac{1}{2}(N + 1)$$

can be attained [8]. It is unknown whether commuting operators can be found whose measurement will yield unbiased estimates $\hat{m}_x$ and $\hat{m}_y$ with variances lying between $\frac{1}{2}(N + \frac{1}{2})$ and $\frac{1}{2}(N + 1)$.

The measurements we need to make on a quantum-mechanical system S for testing hypotheses or estimating parameters will always have to be effected by means of an auxiliary apparatus A, and this apparatus, subject to thermal and quantum fluctuations of its own, will ordinarily introduce additional random uncertainties. Each measurement procedure will have to be analyzed to determine what error costs it entails. Detection theory and estimation theory seek lower bounds on these costs, and in doing so they minimize with respect to the entire class of possible detection or estimation operators that can be applied to the system. The resulting bounds are independent of the time of observation, and they cannot be reduced by using any auxiliary apparatus that initially posseses no information about the state of the system.

References

1. C. W. Helstrom, "Detection theory and quantum mechanics", Infor. and Control, vol. 10, pp. 254-291; March, 1967.

2. C. W. Helstrom, "Detection theory and quantum mechanics, II", Infor. and Control, vol. 13, pp. 156-171; August, 1968.

3. C. W. Helstrom, "The minimum variance of estimates in quantum signal detection", IEEE Trans. on Information Theory, vol. IT-14, pp. 234-242; March, 1968.

4. C. W. Helstrom, J. W. S. Liu, and J. P. Gordon, "Quantum-mechanical communication theory", Proceedings of the IEEE, vol. 58, pp. 1578-1598; October, 1970.

5. H. P. Yuen, R. S. Kennedy, and M. Lax, "On optimal quantum receivers for digital signal detection", Proceedings of the IEEE, vol. 58, pp. 1770-1773; October, 1970.

6. J. P. Gordon and W. H. Louisell, "Simultaneous measurement of noncommuting observables", in Physics of Quantum Electronics, P. L. Kelley, B. Lax, and P. E. Tannenwald, Eds. New York: McGraw-Hill, 1966, pp. 833-840.

7. W. H. Louisell, Radiation and Noise in Quantum Electronics. New York: McGraw-Hill, 1964, p. 225.

8. C. W. Helstrom, "Quantum detection theory", to appear in Progress in Optics, E. Wolf, Ed. Amsterdam: North-Holland Publishing Co.